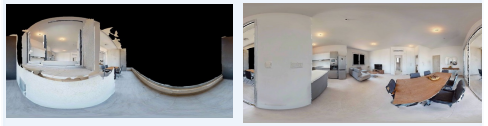
Route Filtering

1. Rendering Quality

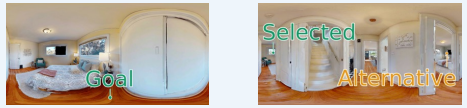


✕ Mesh holes ✓ Intact mesh

2. Route Decoupling

Similar routes → Keep one

3. Replay validation



✓ Goal reached ✓ Choices visible

State sampling

Dense · stride 1 ✕

≈ ≈ Similar views

t
t+1
t+2

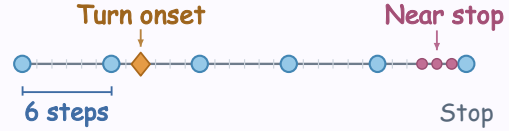
Adjacent overlap: 17 / 18

Sparse · stride 6 ✓

t
t+6
t+12

Adjacent overlap: 12 / 18

Add event states ✓

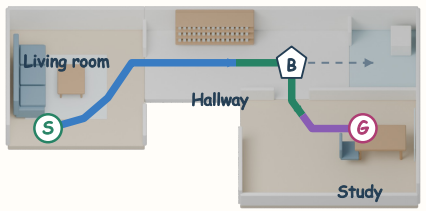


Regular grid + extra event samples

Instruction construction

1 Route segmentation

Split at meaningful route events



Travel | Branching | Arrival

Branching point

2 Local annotation

Choose evidence for each segment type

Travel — Video → VLM → Leave the living room; follow the hallway.

Branching — Video + 8-view compass → VLM → In the hallway, turn right into the study.

Arrival — Video + 8-view compass → VLM → In the study, stop beside the desk.

3 Composition

Join clauses and remove repetition

1 …follow the hallway.
2 In the hallway, turn right into the study.
3 In the study, stop beside the desk.

↓ Remove repeated context; keep route order

Leave the living room and follow the hallway, then turn right into the study and stop beside the desk.

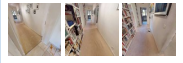
4 Verification

Check the instruction against clean evidence

Observe clean evidence → Compare local instruction

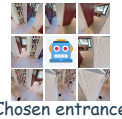
No instruction / route overlays

Motion consistency



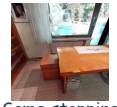
Movement order and turn directions

Choice grounding



Chosen entrance vs. alternatives

Stop grounding



Same stopping area

Mismatch → Local revision

All consistent → accept